

母分散の不偏推定値はなぜ $n-1$ で割られるのか (2) Why divide by $n-1$ when we estimate population variance (2)

野口 智草・宇津木 成介
Chigusa Noguchi, Narisuke Utsuki

要 旨 (Abstract)

標本の分散から母分散を推定する場合、標本の偏差平方和を n で割らずに $n-1$ で割ることが必要であるが、このことを文系の学生に教えることは必ずしも容易ではないことを前報で示した。本稿では、実例を複数挙げることによって説明に説得力を持たせる方法を具体的に示した。

キーワード：(母集団分散) (標本分散) (不偏推定値)

I はじめに

母集団からランダムに n 個の標本を取り出してその値を取り出し、平均値を求めると、これは母集団平均値の最良の推定値となる (標本測定値の総和を n で割った数値は、母平均の不偏推定値である)。このような平均値を多数回算出し、その平均をとると、その平均値は母集団平均値に近づくからである。心理学を専攻する多くの学生は、このことを容易に理解する。しかし、母集団の分散を推定する場合に、偏差平方和を n ではなく、 $n-1$ で割らないと母分散の普遍推定値が得られないことを容易に理解する学生は、非常に少ないと思われる。経験上、多くの学生は、「標本測定値自体の分散を求めるときは n で割り、母集団の分散を推定するときには $n-1$ で割る」という天下り式の説明を受け入れる。しかし少数ではあるが、このような天下り式の説明に納得しない学生も散見される。前報で筆者らは、なぜ n ではなく $n-1$ で割ると不偏推定値が得られるのかについて、統計学の初等教科書がどのように説明しているかを検討した (宇津木・野口, 2016) ⁽¹⁾。

母集団の分散の不偏推定では、なぜ偏差平方和を $n-1$ で割るのを説明するのかという問題 (「 $n-1$ 問題」と書く) について、日本で書かれた教科書 7 点とアメリカ合衆国で書かれた教科書 (邦訳されたものを含む) 6 点の説明を精査した結果、いずれにも厳密な証明や説明は与えられていなかった。検討対象とした書籍の中では、実例を挙げることで学習者の説得を試みた「心理教育統計学 (肥田野・瀬谷・大川, 1961)」の方向性が好ましく思われた。そこで示されている例は、6 つの値をもつ無限母集団から 2 つの測定値を取り出す (サンプルサイズが 2) という例である。しかし、サンプルサイズが 2 の場合に偏差平方和を 2 ではなく 1 で割ると母分散の不偏推定値になるというだけでは、一般に標本の偏差平方和を $n-1$ で割ると母分散の不偏推定値になることを説得するには足りない

⁽¹⁾ 宇津木・野口 (2016) では論文題名中「不偏推定値」としなければならないところを、「不変推定値」と誤記したまま気づかずに公刊した。これは当該論文の第一著者 (本論文の第二著者) の校正の不行き届きに起因する。読者にお詫びを申し上げますとともに、ここに訂正する。

ように思われた。

一方、「心理学のためのデータ解析テクニカルブック」(森・吉田, 1990)は、概略、「母平均 μ がわからないために標本平均値 \bar{X} を代用した。さて \bar{X} は μ を中心に σ^2/n の分散で分布している。従って母分散は、 \bar{X} を中心とする X の分散と、 μ を中心とする \bar{X} の分散の和となる。」という説明によって、標本の偏差平方和を $n-1$ で割ると母分散の推定値となることを示そうとしている。この説明は合理的であるし、学生に対して誠意をもって統計学を教えようとする意図が見えて好ましい。ただ、分散の分解(分散分析)を知らないとこの説明が理解できないと思われること、「 \bar{X} は μ を中心に σ^2/n の分散で分布する」ことが天下り的に使われていることが難点であると思われる。

そこで本稿においては、1) 複数の例を上げ、いずれの場合にも標本の偏差平方和を $n-1$ で割った値が母分散の不偏推定値になることを示すこと、2) 森・吉田(1990)の説明の中心である「標本平均値 \bar{X} は母平均 μ を中心に σ^2/n の分散で分布する」ことを、例を挙げて示すことによって、心理学の学生の疑念を取り除くことを試みた。

II 標本の偏差平方和から母分散を推定する

1. 2値の母集団から標本2個を取り出す

無限母集団があり、その中には1と-1が同数(1対1の割合で)含まれている。例えば、男女が半々の母集団があつて、男子は-1と印刷された札を持ち、女子は+1と印刷された札を持っているとする。ランダムに1名を選び、持っている札の数字を記録する。

母集団には-1と+1が半々の割合で含まれているので、母集団の平均値は0である。母集団の分散 V は、 $V = ((-1)^2 + 1^2)/2$ 、つまり母分散は1である。

さて実際にこの母集団から2つの標本をとってくる。最初に1を取り、2回めにも1を取る場合を(1, 1)と書くことにする。そうすると標本の取り方は、(1, 1) (1, -1) (-1, 1) (-1, -1)の4通りの場合しかない。また、この4通りはすべて同じ確率で生じるだろう。この4通りの場合について、それぞれ、標本平均と標本分散を算出してみる。

(1, 1)の場合、平均値は1、偏差平方和は0、分散は0である。

(1, -1)の場合、平均値は0、偏差平方和は2、分散は1である。

(-1, 1)の場合、平均値は0、偏差平方和は2、分散は1である。

(-1, -1)の場合、平均値は-1、偏差平方和は0、分散は0である。

これらから、標本平均値の平均は0、偏差平方和の平均値は1であり、標本分散の平均値は0.5である。

標本平均値の平均は母集団の平均値と一致している。そこで、標本平均値を、母集団の平均値の最良の推定値(不偏推定値)と呼ぶ。しかし、分散については、標本分散の平均値は母分散と一致していない。ところで、分散を求めるに際して、偏差平方和をサンプルサイズ(ここでは2)で割る代わりに、(サンプルサイズ-1)で割るとどうなるか。偏差平方和の平均値は1であり、(サンプルサイズ-1)は1であるから、この値は1となり、母分散に一致する。しかしこれは、単純な数を使ったために偶然に生じたのかもしれないという疑念が残る。そこで、次に、この母集団から3つの標本を取り出す場合について調べる。

2. 2値の母集団から標本3個を取り出す

標本を3個取り出すのであるから、(1, 1, 1) (1, 1, -1) (1, -1, 1) (-1, 1, 1) (1, -1, -1) (-1, 1, -1) (-1, -1, 1)

$(-1, -1, -1)$ の 8 通りの取り出し方ができる。そこで先ほどと同様にそれぞれの標本について平均値と分散を計算する。 $(1, 1, 1)$ と $(-1, -1, -1)$ の場合には分散は 0 である。また、 $(1, 1, -1)$ $(1, -1, 1)$ $(-1, 1, 1)$ は、いずれも 1 が 2 つ -1 が 1 つであり、平均値も分散も同じ値であり、 $(1, -1, -1)$ $(-1, 1, -1)$ $(-1, -1, 1)$ についても平均値と分散はいずれも等しい。

この 8 通りの場合の平均値は、1 が 1 通り、 $1/3$ が 3 通り、 $-1/3$ が 3 通り、 -1 が 1 通りであるから、合計するとゼロになり、母平均と一致する。

偏差平方和は、 $(1, 1, -1)$ の場合、 $(1-1/3)^2+(1-1/3)^2+(-1-1/3)^2=24/9$ である。 $(-1, -1, 1)$ の場合も $(-1+1/3)^2+(-1+1/3)^2+(1+1/3)^2=24/9$ であるから、偏差平方和は $24/9$ が 6 通りと 0 が 2 通りである。合計は $48/9$ である。これを 8 で割ると偏差平方和の平均値は 6 である。これをサンプルサイズの 3 で割って得られる 2 が標本分散の平均値である。先ほどと同様に、偏差平方和の平均値を (サンプルサイズ-1) で割ると、 $2 \div 2 = 1$ となって、やはり母分散と一致する。

3. 3 値の母集団から標本 2 個を取り出す

次に、 $+1, 0, -1$ の 3 値を均等に持つ母集団から 2 つのサンプルを取り出す場合について考えてみよう。この母集団の平均値は 0、偏差平方和は 2、分散は $2/3$ である。

この場合、 $(1, 1)$ $(1, 0)$ $(1, -1)$ $(0, 1)$ $(0, 0)$ $(0, -1)$ $(-1, 1)$ $(-1, 0)$ $(-1, -1)$ の 9 通りの標本が得られる。平均値は、それぞれ 1, 0.5, 0, 0.5, 0, -0.5, 0, -0.5, -1 であり、平均値の平均値は 0 である。偏差平方和は、それぞれ 0, 0.5, 2, 0.5, 0, 0.5, 2, 0.5, 0 である。合計は 6 であるから、偏差平方和の平均値は $2/3$ 、標本分散の平均値はサンプルサイズの 2 で割って $1/3$ である。この場合も偏差平方和の平均値をサンプルサイズの 2 ではなく、サンプルサイズから 1 引いた値である 1 で割ると、母分散である $2/3$ が得られる。

4. 具体的数字によらない例

要素として a と b を無限かつ均等にもつ母集団を考える。あるいは、 a と b の数字が書かれた 2 枚の札を袋に入れて 1 枚取り出し、値を記録して袋に戻し、再度 1 枚を取り出す手続きを繰り返してもよい。母平均は $(a+b)/2$ である。母分散を求めるためにまず偏差平方を計算する。偏差平方和は $(a-(a+b)/2)^2+(b-(a+b)/2)^2$ である。

これは、 $(a-b)^2/4+(a-b)^2/4$ であるから、偏差平方和は $(a-b)^2/2$ である。従って母分散は $(a-b)^2/4$ である。

1) 標本サイズが 2 の場合

この母集団から要素 1 つを取り出して値を記録してもとに戻す。要素 2 つを取り出すごとに ($n=2$) その分散を計算する。 $n=2$ のとき、取り出せる組み合わせは、 aa, ab, ba, bb の 4 通りである。 aa と bb のときは偏差平方和が 0 であるから、分散は 0 である。取り出した要素が ab と ba のとき、偏差平方和は $(a-(a+b)/2)^2+(b-(a+b)/2)^2$ つまり $(a-b)^2/4+(a-b)^2/4$ である。標本分散は偏差平方和を 2 で割った値であるから、 $(a-b)^2/4$ である。結局、 aa, ab, ba, bb の標本分散はそれぞれ、0, $(a-b)^2/4, (a-b)^2/4, 0$ である。サンプルをとって分散を算出することを繰り返すと、この 4 つの場合が均等に現れる。この 4 つの分散の合計は $(a-b)^2/2$ である。この平均 (標本分散の期待値) は、 $(a-b)^2/2$ の 4 分の 1 であるから、 $(a-b)^2/8$ である。これは、たしかに定義による母集団の分散の $(n-1)/n$ 、つまり 2 分の 1 になっている。

2) 標本サイズが 3 の場合

つぎに 3 枚を取り出す場合を考える。 $n=3$ のとき取り出せる組み合わせは、 $aaa, aab, aba, abb, baa, bab, bba, bbb$

の8通りである(2の3乗通り)。すべてが a または b の場合がそれぞれ1通りある。 a が2個で b が1個の場合、 a が1個で b が2個の場合はそれぞれ3通りある。標本が aaa と bbb の場合、分散は0である。次に a が2個 b が1個の場合、その偏差平方和は、

$$(a-(a+a+b)/3)^2 + (a-(a+a+b)/3)^2 + (b-(a+a+b)/3)^2 \text{ である。}$$

整理すると、 $(a-b)^2/9 + (a-b)^2/9 + 4(a-b)^2/9$ であり、これは $6(a-b)^2/9$ である。標本の分散は偏差平方和を $n=3$ で割った値、 $2(a-b)^2/9$ である。次に、 a が1個、 b が2個の場合の偏差平方和は、同様に $6(a-b)^2/9$ であるから標本の分散は、 $2(a-b)^2/9$ である。

標本の分散の合計は $12(a-b)^2/9$ となり、分散の平均値は $12(a-b)^2/9/8$ つまり、 $(a-b)^2/6$ である。これを $n-1$ で割って n 倍すれば、母分散 $(a-b)^2/4$ が得られる。

3) 標本サイズが4の場合

つぎに4枚を取り出す場合を考える。取り出し方は16通りあり、4枚ともすべて a の場合が1通り(この場合、分散はゼロ)、 a 3つと b 1つの場合が4通り、 a 2つと b 2つの場合が6通り、 a 1つと b 3つの場合が4通り、4枚ともすべて b の場合が1通りである(この場合、分散はゼロ)。

a 3つと b 1つの場合、偏差平方和は

$$(a-(3a+b)/4)^2 + (a-(3a+b)/4)^2 + (a-(3a+b)/4)^2 + (b-(3a+b)/4)^2 \text{ である。}$$

これは $3(a-b)^2/16 + 9(a-b)^2/16$ であるから、結局 $12(a-b)^2/16$ である。 $n=4$ で割って標本分散 $3(a-b)^2/16$ を得る。

a 2つと b 2つの場合、偏差平方和は $2(a-(2a+2b)/4)^2 + 2(b-(2a+2b)/4)^2$

これは、 $2(2(a-b)/4)^2 + 2(2(a-b)/4)^2$ であるから、結局 $(a-b)^2$ である。 $n=4$ で割って、標本分散 $(a-b)^2/4$ を得る。

a 1つと b 3つの場合は a 3つと b 1つの場合と同じ $12(a-b)^2/16$ である。

標本分散の期待値は、16通りの場合の分散の合計、つまり $24(a-b)^2/16 + 6(a-b)^2/4$ である。これを整理すると、 $24(a-b)^2/16 + 24(a-b)^2/16$ となり、最終的に、 $3(a-b)^2$ を得る。

これを16で割った値、つまり $3(a-b)^2/16$ が標本分散の平均値である。この場合も $n-1$ で割って n 倍すれば(4/3を掛ければ)、母分散 $(a-b)^2/4$ が得られる。

以上をまとめると、母分散 $(a-b)^2/4$ を V とすると

$$n=2 \text{ のときの標本分散の平均値は } (a-b)^2/8 = 1V/2$$

$$n=3 \text{ のときの標本分散の平均値は } (a-b)^2/6 = 2V/3$$

$$n=4 \text{ のときの標本分散の平均値は } 3(a-b)^2/16 = 3V/4$$

であり、これらはいずれも母分散 V に $(n-1)/n$ を掛けた値になっている。

逆に、標本分散平均値に $n/(n-1)$ を掛ければ、それが母分散となる。

もちろん、 $n=2, 3, 4$ の時にこのルールが成立するからといって任意の n について同じルールが成立すると数学的に主張できるわけではない。しかし、学生の教育に関するかぎり、統計学の教科書に書いてある通り、任意のサンプルサイズ n についても、標本分散の期待値は母分散の $(n-1)/n$ であり、従って標本の偏差平方和を n で割る代わりに $(n-1)$ で割ると母分散の最も良い推定値(不偏推定値)が得られると、学生に教えることができるだろう。経験上、文系の学生は、天下り式に $(n-1)$ で割ることになっていると言われると反発を感じるが、任意の n について成り立つことを厳密に示さなくても、幾つかの実例でルールが成立することを示せば、納得するようになる。

5) まとめ

m 個の値を均等にもつ母集団から n 個の標本をとる場合の数は m の n 乗であり、それぞれの標本について平均値、偏差平方和、分散を計算し、全標本の平均をとってみると、標本の偏差平方和平均値をサンプルサイズ n でなく、 $n-1$ で割った値が母分散に一致することがわかる。これは、エクセルを使うと簡単に可視化できる。本稿の末尾に例を挙げる (付表)。

なお、各要素が均等でない場合、例えば 1 と 2 からなる母集団であるが、1 は 2 の 2 倍の頻度で現れるという場合、これは、袋の中に 1 の札が 2 枚、2 の札が 1 枚入っていることと同じであり、この場合には 3 個の値を均等に持つ母集団であると考えればよいことを示せばよいであろう。

Ⅲ 標本平均値の分布について

標本分散の平均値 (期待値) と母集団の分散との間に「母分散 $\times(n-1)/n$ =標本分散の平均値」という一定の関係がなぜ成立するのか。 $(n-1)/n$ は $1-1/n$ であるから、実は「」内は、「母分散-母分散 $/n$ =標本の分散の平均値」と書き直すことができる。そして、母分散を n で割った値は、学生が t 検定について学習するときに「サンプルの平均値は母平均を中心として母分散の $1/n$ の分散で分布する」ことを学べば、「なるほど」と了解するかもしれない。以下においては、標本平均値が、母平均を平均値として母分散 $/n$ の分散で分布すること (平均値が μ で分散が V である母集団から得られる n 個の要素からなる標本の平均値 \bar{X} は、母平均 μ を中心に V/n の分散で分布する) を示す。

要素として a と b を無限かつ均等にもつ母集団を考える。あるいは、 a と b の 2 枚の札を袋に入れて 1 枚取り出し、値を記録して袋に戻し、再度 1 枚を取り出す手続きを繰り返してもよい。すでに示したように、母平均は $(a+b)/2$ である。母分散は $(a-b)^2/4$ である。次に、標本平均と母平均の間の偏差平方和を求めることにする。

1. 標本サイズが 2 の場合

母集団から 2 つの要素をとって標本とし、その平均値 (標本平均) を求める。次に、標本平均から母平均を引き、その 2 乗を求める。 $n=2$ であるから、 aa, ab, ba, bb の 4 通りの標本がある。 aa の場合の平均値は a である。 a から母平均 $(a+b)/2$ を引いて、2 乗し、偏差平方 $(a-(a+b)/2)^2$ を得る。これは $(a-b)^2/4$ である。 bb の場合の平均値 b についても同様に計算をする。これは $(b-(a+b)/2)^2$ であるから、やはり $(a-b)^2/4$ である。

要素が ab, ba の場合には平均値は $(a+b)/2$ である。これは母平均と同じであるから、偏差はゼロである。このすべての標本 (4 通り) の平均値の母平均からの偏差平方和は、 $(a-b)^2/4+(a-b)^2/4+0+0$ となる。この平均値は、 $(a-b)^2/8$ である。これが、標本平均値の分散である。上記Ⅱ. 4 から、このような標本の分散の平均値は $(a-b)^2/8$ である。標本の分散の平均値と標本平均値の (母平均に対する) 分散の和は、 $(a-b)^2/8+(a-b)^2/8$ であり、母分散 $(a-b)^2/4$ に一致する。

2. 標本サイズが 3 の場合

母集団から 3 つの要素をとって標本とし、その平均値 (標本平均値) を求める。次に、標本平均値から母平均を引き、その 2 乗を求める。 $n=3$ であるから、 $aaa, aab, aba, abb, baa, bab, bba, bbb$ の 8 通りの標本がとれる。すべての要素が a の場合が 1 通り、 a が 2 つで b が 1 つの場合が 3 通り、 a が 1 つで b が 2 つの場合が 3 通り、すべ

ての要素が b である場合が 1 通りである。偏差平方和は、以下の項の合計となる。

$$(a-(a+b)/2)^2$$

$$3 \times ((2a+b)/3-(a+b)/2)^2$$

$$3 \times ((a+2b)/3-(a+b)/2)^2$$

$$(b-(a+b)/2)^2$$

このうち第 1 項と第 4 項は既出のように $(a-b)^2/4$ である。残りの第 2 項と第 3 項について展開する。第 2 項について詳しく述べると、

$$3 \times \left[\frac{4a^2+4ab+b^2}{9} - \frac{(2a+b)(a+b)}{3} + \frac{(a+b)^2}{4} \right]$$

$$= 3 \times \left[\frac{1}{9} (4a^2+4ab+b^2-6a^2-9ab-3b^2) + \frac{1}{4} (a^2+2ab+b^2) \right]$$

$$= 3 \times \left[\frac{1}{36} (-8a^2-20ab-8b^2) + \frac{1}{36} (9a^2+18ab+9b^2) \right]$$

$$= 3 \times \frac{1}{36} (a^2-2ab+b^2)$$

$$= (a-b)^2/12$$

第 3 項は a と b を置き換えるだけであるから、 $(b-a)^2/12$ であり、これは $(a-b)^2/12$ と等しい。第 1 項から第 4 項までの和である偏差平方和は、 $2(a-b)^2/4 + 2(a-b)^2/12$ つまり $2(a-b)^2/3$ になる。

標本の数 n は 8 であるから、母平均からの偏差平方和を 8 で割った値、つまり標本平均値の分散は、 $(b-a)^2/12$ となる。上記 II. 4 から $n=3$ の場合の標本分散の平均値は $(a-b)^2/6$ である。標本分散の平均値である $(a-b)^2/6$ と標本平均値の分散 $(b-a)^2/12$ とを加えると $(a-b)^2/4$ になる。これは母分散にほかならない。標本の分散の平均値と標本平均値の（母平均に対する）分散の合計は、ここでも母分散に一致する。

3. 標本サイズが 4 の場合

母集団から要素を 4 つとる $n=4$ のとき、標本の取り方は 16 通りある。途中は省略するが、標本平均値の母平均に対する偏差平方和は $16(a-b)^2/16$ となる。これを 16 で割れば、 $(a-b)^2/16$ が得られる。II. 4 から $n=4$ のときの標本分散の平均値は $3(a-b)^2/16$ である。標本分散の平均値と標本平均の母平均に対する分散の和は、ここでも $(a-b)^2/4$ となり、母分散に一致する。

4. この項のまとめ

標本の分散は、標本平均値が母平均から離れている程度を考慮していないために、母分散より必ず小さくなるが、 $n/(n-1)$ を掛けることによって母分散の推定値として適切（不偏）になることがわかる。ここで母分散 V を σ^2 、標本分散を s^2 、標本の要素数（サンプルサイズ）を n とすると、一般に、 $\sigma^2 = s^2 + \sigma^2/n$ 、あるいは、 $s^2 = \sigma^2(n-1)/n$ と書くことができる。もし、このことを先に学生に理解させることができるならば、標本から母分散を推定する場合になぜ $n-1$ で偏差平方和を割らねばならないのか、その理由は明白となるだろう。

Ⅲ. 全体のまとめ

筆者らは、数学的に厳密な説明を行う知識も能力も持たないが、主として心理学の学生が初歩の統計学を学ぶ場合に、つまづく可能性のある、母分散の不偏推定値算出について、複数の例を示すことによって学生を説得する方法を案出しようと試みた。筆者らは、数学的知識に欠ける文系の大学生に「なるほど」と感じさせ、教員の教育を疑わせない程度の説明方法があれば、数学的厳密性はかなりの程度犠牲にしてもよいのではないかと考えている。統計学教育の立場から批判はあるだろうが、「昔、 $n-1$ で割るものだと言ったが、いまだになぜそうなのかわからない」という学生に、「なるほど $n-1$ で割ると母分散の推定値として最良なのだ」と納得してもらうことができれば幸いである。

文献

肥田野・瀬谷・大川 (1961). 心理教育統計学 培風館

森 敏昭・吉田寿夫 (1990). 心理学のためのデータ解析テクニカルブック 北大路書房

宇津木成介・野口智草 (2016). 母分散の不偏推定値はなぜ $n-1$ で割られるのか 奈良学園大学紀要第5集 187-195.

付表 要素として1と-1を均等につづ母集団から得られる標本の平均値と分散
(標本サイズが $n=2, 3, 4, 5$ の場合)

n=2				n=3				n=4				n=5			
平均値		分散		平均値		分散		平均値		分散		平均値		分散	
1	1	1	0	1	1	1	1.00	0.00	1	1	1	1	1	1.00	0.00
1	-1	0	1	1	1	-1	0.33	0.89	1	1	1	-1	-1	0.60	0.64
-1	1	0	1	1	-1	1	0.33	0.89	1	1	-1	1	-1	0.60	0.64
-1	-1	-1	0	1	-1	-1	-0.33	0.89	1	1	-1	-1	-1	0.20	0.96
平均		0	0.5	-1	1	1	0.33	0.89	1	-1	1	1	1	0.60	0.64
				-1	1	-1	-0.33	0.89	1	-1	-1	-1	-1	0.20	0.96
				-1	1	1	0.33	0.89	1	-1	1	1	1	0.20	0.96
				-1	1	-1	-0.33	0.89	1	-1	-1	-1	-1	0.20	0.96
				-1	-1	1	-0.33	0.89	1	-1	-1	1	-1	-0.20	0.96
				-1	-1	-1	-1.00	0.00	1	-1	-1	-1	-1	-0.20	0.96
				平均		0.00	0.67	-1	1	1	1	1	1	0.60	0.64
								-1	1	1	-1	1	-1	0.20	0.96
								-1	1	-1	1	1	1	0.20	0.96
								-1	1	-1	-1	-1	-1	-0.20	0.96
								-1	-1	1	1	1	1	0.20	0.96
								-1	-1	1	-1	-1	-1	-0.20	0.96
								-1	-1	-1	1	1	1	0.20	0.96
								-1	-1	-1	-1	-1	-1	-0.20	0.96
								-1	-1	-1	-1	1	1	-0.20	0.96
								-1	-1	-1	1	-1	-1	-0.20	0.96
								-1	-1	-1	-1	1	1	-0.20	0.96
								-1	-1	-1	-1	-1	-1	-0.60	0.64
								-1	-1	-1	-1	1	1	0.20	0.96
								-1	-1	-1	1	1	1	-0.20	0.96
								-1	-1	-1	-1	-1	-1	-0.20	0.96
								-1	-1	-1	-1	1	1	-0.20	0.96
								-1	-1	-1	-1	-1	-1	-0.60	0.64
								-1	-1	-1	-1	1	1	-0.20	0.96
								-1	-1	-1	-1	-1	-1	-0.60	0.64
								-1	-1	-1	-1	-1	1	-0.60	0.64
								-1	-1	-1	-1	-1	-1	-1.00	0.00
								平均		0.00	0.75			0.00	0.80

標本サイズ n が 2, 3, 4, 5 の場合の、とりうる標本とその平均と分散を示す。それぞれの標本サイズにおける「分散」は標本分散であり、エクセルの VARP 関数によって算出したものである。最下段は、AVERAGE 関数によって算出した標本分散の平均値。母集団の分散は 1 である。標本分散の平均値はそれぞれ、母分散の $n-1/n$ になっている。